To: Peters, Emily (MPCA)[Emily.Peters@state.mn.us]
Cc: russerick@chartermi.net[russerick@chartermi.net]
From: Erickson, Russell
Sent: Thur 2/13/2014 10:41:00 PM
Subject: RE: 4-parameter logistic model for pooled hydroponic data

I don't know yet when I will be in tomorrow. Will be out-of-town for most of the day, but not yet sure when. Would 8-9 work out okay? Probably can delay leaving until then.

Russell Erickson

Mid-Continent Ecology Division

National Health and Environmental Effects Research Laboratory

U.S. Environmental Protection Agency

6201 Congdon Boulevard

Duluth, MN, USA 55804

Phone: 218-529-5157

Fax: 218-529-5003

Email: erickson.russell@epa.gov

From: Peters, Emily (MPCA) [mailto:Emily.Peters@state.mn.us]
Sent: Thursday, February 13, 2014 4:23 PM
To: Erickson, Russell
Subject: RE: 4-parameter logistic model for pooled hydroponic data

Russ,

Are you available tomorrow (Friday) to chat again? I have enough complicated questions building that it would be more efficient for me to talk on the phone.

Also, I'm pretty sure Ed and Phil want EC20s. They've worked out the justification for this already. I'm not sure I can get confidence limits out of R on an EC20 estimate. I need to do some investigating still. In any case, yes, it would be amazingly helpful if you can come up with an equation I can use instead.

Thanks again for all your help. FYI – the goal is to have these EC20 values (with or without confidence intervals) by tomorrow (ahhh).

-Emily

From: Erickson, Russell [mailto:Erickson.Russell@epa.gov]
Sent: Thursday, February 13, 2014 3:39 PM
To: Peters, Emily (MPCA)
Subject: RE: 4-parameter logistic model for pooled hydroponic data

Emily:

(1) Yes, this is the model, given that you chose to keep 1/D as the exponent. However, this D will still differ from your original equation by a factor that accounts for having the natural exponent and base 10 logarithms in the original equation.

(2) Computing the EC20 will require just a simple rearrangement. However, this depends on your definition of what 20% means. Your formula would have it mean 20% of the way to the lower plateau (negative growth). I would suggest that it instead be 20% of the way to zero growth, so that y_EC20=0.8*A.

However, whichever way you define 20%, this just gets you the EC20, not its confidence limits. One thing that TRAP provides (and a main reason I wrote it) is confidence limits on any EC. Your regression analysis does give you the confidence limits (or at least the standard error) on the parameters, which include the EC50 – can it also do standard errors or confidence limits on other points of the regression line?

If not, the way TRAP does this is to have the parameter be C'=EC20 rather than C=EC50. If B=0, this is rather simple – the denominator becomes (1+0.25*(x/C')^(1/D)) so it =1.25 when x=C' (rather than =2 when x=C). I am pretty sure that is correct, but it is a lot more complicated it B isn't 0. I can probably come up with an equation you can plug into your R analysis, but would need some time to work on it.

An alternative is to use the EC50 since it isn't that much different than the EC20. This is a more robust and certain endpoint, and since the main point of this experiment is to establish causality and the general concentration range at which effects occur. After all, the regression analysis just gives the confidence limits assuming that the concentrations are known without error – but since the concentrations themselves are so variable and uncertain, are the confidence limits on the EC20, or the distinction between the EC20 and EC50, actually that important?

Russell Erickson

Mid-Continent Ecology Division

National Health and Environmental Effects Research Laboratory

U.S. Environmental Protection Agency

6201 Congdon Boulevard

Duluth, MN, USA 55804

Phone: 218-529-5157

Fax: 218-529-5003

Email: erickson.russell@epa.gov

From: Peters, Emily (MPCA) [mailto:Emily.Peters@state.mn.us]
Sent: Thursday, February 13, 2014 2:58 PM
To: Erickson, Russell
Subject: RE: 4-parameter logistic model for pooled hydroponic data

Thanks, Russ. I have a couple more questions for you:

1. Can you verify that I have the model you're suggesting correct?

a. Model: y = A+(B-A)/(1+(x/C)^(1/D))

b. Y = growth rate constant

c. X = mean sulfide concentration (not log10(mean sulfide concentration))

2. I'm confused how to calculate EC20, once the model parameters are identified. I read the explanation in the TRAP model, but it's not entirely clear to me. I should just be able to rearrange the equation and solve for X given y_EC20 = A - (A-B)*0.2. Right?

Thanks.

-Emily

From: Erickson, Russell [mailto:Erickson.Russell@epa.gov]
Sent: Thursday, February 13, 2014 2:50 PM
To: Peters, Emily (MPCA)
Subject: RE: 4-parameter logistic model for pooled hydroponic data

Emily:

Looks good to me – this is the approach I was describing.

To be simpler and more clear, the denominator could be 1+(x/C)^D'. If done this way, x and C are in concentration units, not log concentration (although I would still plot things on a log scale). D' is the inverse of your D and also differs by a factor of ln(10) because the current expression mixes the natural exponent with base 10 logs (this awkward mixture is one reason to do this simplification – TRAP uses the samemixture, but for reasons not relevant here).

Russ

Russell Erickson

Mid-Continent Ecology Division

National Health and Environmental Effects Research Laboratory

U.S. Environmental Protection Agency

6201 Congdon Boulevard

Duluth, MN, USA 55804

Phone: 218-529-5157

Fax: 218-529-5003

Email: erickson.russell@epa.gov

From: Peters, Emily (MPCA) [mailto:Emily.Peters@state.mn.us]
Sent: Thursday, February 13, 2014 12:41 PM
To: Erickson, Russell
Subject: 4-parameter logistic model for pooled hydroponic data

Russ,

Thanks again for the REALLY helpful advice. I'm attaching the pooled hydroponic data from the 3 tests. Let me know if the headers are not clear. These are the data I'm fitting in R using the following 4-parameter logistic model:

y = A+(B-A)/(1+exp((C-x)/D))

- A = left side horizontal asymptote (max response)
- B = right side horizontal asymptote (min response)
- C = x value (concentration) at the inflection point of the curve. EC50.
- D = slope parameter, indicating the speed with which the curve rises between min and max.

I ran a quick analysis in R that fixed parameters B, C, and D and allowed A to vary among the 3 tests. I created dummy variables (rangefinder, d1, d2) to distinguish the 3 tests in the model (value = 0 or 1). Below is the model and summary statistics. Does this look right to you? Is this how you would set up the model equation?

Formula: length_rate_constant ~ A1*rangefinder + A2*d1 + A3*d2 + (B - A1*rangefinder -

A2*d1 - A3*d2)/(1 + exp((C -log_meansulfide_ugL)/D))

Parameters:

Estimate Std. Error t value Pr(>|t|)

A1 0.084741 0.005650 14.997 < 2e-16 ***

A2 0.065548 0.005407 12.123 8.38e-15 ***

A3 0.059435 0.005416 10.974 1.73e-13 ***

B -0.012143 0.008701 -1.396 0.17073

C 2.655272 0.081060 32.757 < 2e-16 ***

D 0.198887 0.060079 3.310 0.00201 **

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01437 on 39 degrees of freedom

Number of iterations to convergence: 11

Achieved convergence tolerance: 3.923e-06

Thanks much!

-Emily

Emily B Peters, Ph.D.

Data Analyst

Data Services Section

Minnesota Pollution Control Agency

520 Lafayette Rd N, St. Paul, MN 55155

651-757-2860

Emily.Peters@state.mn.us